JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 41, 373–383 (1973)

Solutions for Infinite-Matrix Differential Equations

LEONARD SHAW

Department of Electrical Engineering, Polytechnic Institute of Brooklyn, Brooklyn, New York 11201

Submitted by Richard Bellman

1. INTRODUCTION

Sufficient conditions will be given for the existence, uniqueness, and finite-dimensional approximation of solutions to the infinite dimensional system

$$\dot{\mathbf{x}}(t) = A(t)\,\mathbf{x}(t); \qquad \mathbf{x}(t_0) = \mathbf{c} \tag{1.1}$$

in which $\mathbf{x}(t)$ has elements $x_i(t)$: $i = 1, 2,...$ and $A(t)$ has elements $a_{ij}(t)$; $i, j = 1, 2,...$.

The results presented here cover cases which are not included in the conditions given by Bellman in [1]. The present conditions require uniform finiteness, *separately*, of row sums and column sums

$$\sum_{j=1}^{\infty} | a_{ij}(t)| \qquad \text{and} \qquad \sum_{i=1}^{\infty} | a_{ij}(t)| , \tag{1.2}$$

whereas [1] required finiteness of matrix norms of the form

$$N_p(A) = \left[\sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} | a_{ij} |^{p'}\right)^{p-1}\right]^{1/p}$$
$$p > 1; \qquad \left(\frac{1}{p} + \frac{1}{p'}\right) = 1 \tag{1.3}$$

or

$$N_1(A) = \sum_{i,j=1}^{\infty} | a_{ij} | . \tag{1.4}$$

The proofs here parallel those of [1], but inequality properties of matrix norms must be replaced by other inequalities which are justified through uniform convergence arguments and consequent interchanges of the orders of limiting operations. This alternate approach led in [2] to existence and finite-dimensional approximations of solutions to systems like (1.1) having *time-invariant* A matrices with finite row-sums. The additional column-sum condition introduced in [2] for the uniqueness proof is used here to get existence and finite-dimensional approximation of solutions for the case when $A(t)$ is a *time-varying* matrix.

2. Summary of Results

The system in (1.1) will be considered over finite time intervals, so there is no loss in generality in using the interval $\mathscr{T} = \{t : 0 \leqslant t \leqslant T\}$ with $t_0 = 0$. Acceptable solutions will be limited to infinite dimensional vector time functions in a class $\mathscr{C}$ defined by the following three conditions which must hold for all $t \in \mathscr{T}$:

$$\begin{aligned} &\text{(i)} \quad x_i(t) \text{ absolutely continuous,} \\ &\text{(ii)} \quad \| \mathbf{x}(t) \| = \sum_{i=1}^{\infty} | x_i(t) | \leqslant \gamma < \infty, \\ &\text{(iii)} \quad \| \mathbf{x}(t) \| \text{ continuous.} \end{aligned} \tag{2.1}$$

Theorems 7.12 and 7.13 of [3] allow an equivalent definition of $\mathscr{C}$ using (i), (ii) and

(iii′) The sum in (2.1-ii) converges uniformly for $t \in \mathscr{T}$.

Differentiation of the infinite vector $\mathbf{x}(t)$ is defined by specifying the j-th element of $\dot{\mathbf{x}}(t)$ as

$$\dot{x}_j(t) = (d/dt)\, x_j(t). \tag{2.2}$$

Several other theorems on uniform convergence, contained in [3] will be used below.

The vector norm defined in (2.1-ii) will be used throughout. The results are summarized in the following two theorems.

Theorem 1. *There exists a unique vector in* $\mathscr{C}$ *which is a solution to* (1.1) *for* $t \in \mathscr{T}$ *if each* $a_{ij}(t)$ *is measurable and*

$$\begin{aligned} &\text{(a)} \quad \| \mathbf{c} \| \leqslant M < \infty, \\ &\text{(b)} \quad \sum_{j=1}^{\infty} | a_{ij}(t) | \leqslant \alpha < \infty \qquad \textit{all } i,\ t \in \mathscr{T}, \\ &\text{(c)} \quad \sum_{i=1}^{\infty} | a_{ij}(t) | \leqslant \beta < \infty \qquad \textit{all } j,\ t \in \mathscr{T}, \end{aligned} \tag{2.3}$$

with the convergence of the sum in (c) *being uniform over* $t \in \mathscr{T}$ *for each* j.

A collection of finite-dimensional systems

$$\dot{\mathbf{x}}^{(n)}(t) = A^{(n)}(t); \qquad \mathbf{x}^{(n)}(0) = \mathbf{c}^{(n)} \tag{2.4}$$

can be defined by truncating the original system of (1.1). It is convenient to express these finite dimensional systems as infinite dimensional ones with all but a finite number of elements being zero. In particular, the n-dimensional system is defined by

$$\begin{aligned} c_i^{(n)} &= c_i; & i &= 1, 2,..., n \\ &= 0; & i &> n, \\ a_{ij}^{(n)} &= a_{ij}; & i, j &= 1, 2,..., n \\ &= 0; & i &\text{ or } j > n. \end{aligned} \tag{2.5}$$

(Note that $a_{ij}^{(n)}$ had a different meaning in [2].)

Theorem 2. *The conditions of Theorem* 1 *assure that unique solutions* $\mathbf{x}^{(n)}$ *exist for the systems defined in* (2.4) *and* (2.5) *and that the sequence* $\mathbf{x}^{(1)}, \mathbf{x}^{(2)},...$ *converges uniformly over* $\mathscr{T}$ *to the solution of* (1.1).

3. Proofs of Theorems

The proofs given here follow the pattern of similar ones in [1], but the relations which existed there due to the properties of matrix norms must be justified here by uniform convergence conditions and interchanges of limiting operations. These proofs work with the equivalent integral equation form of (1.1):

$$\mathbf{x}(t) = \mathbf{c} + \int_0^t A(\tau)\,\mathbf{x}(\tau)\,d\tau. \tag{3.1}$$

The existence part of Theorem 1 uses a sequence of functions $\mathbf{x}[1],...$ which is defined (with the time-dependence suppressed in the notation) by

$$\begin{aligned} \mathbf{x}[0] &= \mathbf{c}, \\ \mathbf{x}[k] &= \mathbf{c} + \int_0^t A\mathbf{x}[k-1]\,d\tau; \qquad k \geqslant 1. \end{aligned} \tag{3.2}$$

The proof will first establish that $\mathbf{x}[k] \in \mathscr{C}$ for every k, and then that $\mathbf{x}[k]$ converges uniformly to a solution of (1.1).

Clearly $\mathbf{x}[0] \in \mathscr{C}$. Induction on the elements $x_i[k]$ will show that all $\mathbf{x}[k] \in \mathscr{C}$. From (3.2),

$$| x_i[k]| \leqslant | c_i | + \int_0^t \sum_j | a_{ij} | \, | x_j[k-1]| \, d\tau. \tag{3.3}$$

(Summations without limits should be understood to run from 1 to ∞.) With $\mathbf{x}[k-1] \in \mathscr{C}$ and using the assumed uniform boundedness of $|a_{ij}|$, the sum in (3.3) converges uniformly, justifying term-by-term integration. Thus, moving the integration inside the j-sum, and summing over i produces

$$\|\mathbf{x}[k]\| \leqslant \|\mathbf{c}\| + \sum_j \sum_i \int_0^t |a_{ij}| \, |x_j[k-1]| \, d\tau, \tag{3.4}$$

where the order of sums of nonnegative terms also has been interchanged. The assumed uniform convergence in (2.3-c) and the uniform boundedness of each element of a vector $\mathbf{x}[k-1] \in \mathscr{C}$ allow reversal of the order of integration and i-summation, with the result

$$\|\mathbf{x}[k]\| \leqslant \|\mathbf{c}\| + \sum_j \int_0^t \beta \, |x_j[k-1]| \, d\tau. \tag{3.5}$$

Use again of the assumed uniform convergence of $\|\mathbf{x}[k-1]\|$ leads to

$$\|\mathbf{x}[k]\| \leqslant \|\mathbf{c}\| + \beta \int_0^t \|\mathbf{x}[k-1]\| \, d\tau. \tag{3.6}$$

Iterative use of (3.6) verifies the common bound

$$\|\mathbf{x}[k]\| \leqslant \|\mathbf{c}\| \, e^{\beta t}; \qquad k = 0, 1, 2, \ldots. \tag{3.7}$$

This relation is clearly true for $k = 0$. The iterations show that if $\mathbf{x}[0], \mathbf{x}[1], \ldots, \mathbf{x}[k-1]$ satisfy (3.7), then so does $\mathbf{x}[k]$. This establishes that all functions in the sequence defined by (3.2) satisfy property (2.1-ii) in the definition of $\mathscr{C}$, with

$$\gamma = \|\mathbf{c}\| \, e^{\beta T}. \tag{3.8}$$

With regard to the other properties of members of $\mathscr{C}$, the continuity of $x_i[k]$ for all i and k follows from their definition in (3.2) in terms of integrals with bounded integrands. Finally, (2.2) holds for $\mathbf{x}[k]$ if $\mathbf{x}[k-1] \in \mathscr{C}$, for the following reasons. Increasing the right side of (3.3) by letting $t = T$ yields

$$|x_i[k]| \leqslant |c_i| + \sum_j \int_0^T |a_{ij}| \, |x_j[k-1]| \, d\tau. \tag{3.9}$$

Furthermore, the partial sums of the right sides of (3.9) increase monotonically toward the γ in (3.8). Thus $\|\mathbf{x}[k]\|$ converges uniformly since its terms are, respectively, bounded in magnitude by those of a sum which converges uniformly.

Now that it has been shown that $\mathbf{x}[k] \in \mathscr{C}$ for every k, it remains to show the convergence of this sequence of vectors to a solution of (1.1).

Arguments parallel to the preceding ones lead to

$$\begin{aligned} \| \mathbf{x}[1] - \mathbf{x}[0] \| &\leqslant \beta \int_0^t \| \mathbf{x}[0] \| \, d\tau, \\ \| \mathbf{x}[k+1] - \mathbf{x}[k] \| &\leqslant \beta \int_0^t \| \mathbf{x}[k] - \mathbf{x}[k-1] \| \, d\tau; \qquad n \geqslant 1. \end{aligned} \tag{3.10}$$

These relations are a special case of Bellman's (1.17) and (1.18) in [1], with β replacing $N_1(A)$ and $p = 1$. It follows that the remainder of the *existence* proof in that reference carries over to the present situation, with the corresponding substitutions in all steps.

The *uniqueness* part of Theorem 1 is proved by showing that every pair of admissible solutions to (1.1) must have a difference whose norm is zero. If $\mathbf{x}$ and $\mathbf{y}$ are two such solutions of (1.1) and (3.1), then

$$| x_i - y_i | \leqslant \int_0^t \sum_j | a_{ij} | \, | x_j - y_j | \, d\tau. \tag{3.11}$$

Uniform convergence of $\| \mathbf{x} \|$ and $\| \mathbf{y} \|$, and the uniform boundedness of $| a_{ij} |$ allow reversal of the order of summation and integration. Doing so, and summing over i yields

$$\| \mathbf{x} - \mathbf{y} \| \leqslant \sum_j \sum_i \int_0^t | a_{ij} | \, | x_j - y_j | \, d\tau, \tag{3.12}$$

where the order of sums of nonnegative terms also has been reversed. The uniform convergence of (2.3-c) and the uniform boundedness of $| x_j |$ and $| y_j |$ permit summation over i before integration, with the result

$$\| \mathbf{x} - \mathbf{y} \| \leqslant \sum_j \int_0^t \beta \, | x_j - y_j | \, d\tau. \tag{3.13}$$

One more justifiable reordering of limiting operations yields

$$\| \mathbf{x} - \mathbf{y} \| \leqslant \beta \int_0^t \| \mathbf{x} - \mathbf{y} \| \, d\tau, \tag{3.14}$$

which is the same as (3.42) in [2] or (after the substitution of β for $N_1(A)$) the same as (1.23) in [1]. The remainder of either of those uniqueness proofs will complete the present uniqueness proof.

Theorem 1 has the useful byproduct of proving the existence and uniqueness of the solutions (in $\mathscr{C}$) to the finite-dimensional truncations of (1.1), which were defined in (2.4). The bulk of the proof of *Theorem* 2 will establish that for every $\rho > 0$, sufficiently large m and n assure that

$$\| \mathbf{x}^{(n)}(t) - \mathbf{x}^{(m)}(t) \| \leqslant \rho; \qquad \text{all } t \in \mathscr{T}. \tag{3.15}$$

This Cauchy conditions implies that the $\mathbf{x}^{(n)}$ sequence converges uniformly to an $\mathbf{x}^*$ in $\mathscr{C}$. Such convergence, along with the obvious limits, as $n \to \infty$,

$$\mathbf{c}^{(n)} \to \mathbf{c}, \qquad A^{(n)} \to A \tag{3.16}$$

shows that the limit of the integral equation form of (2.4),

$$\mathbf{x}^{(n)} = \mathbf{c}^{(n)} + \int_0^t A^{(n)} \mathbf{x}^{(n)} \, d\tau, \tag{3.17}$$

is

$$\mathbf{x}^* = \mathbf{c} + \int_0^t A \mathbf{x}^* \, d\tau, \tag{3.18}$$

so that the limiting $\mathbf{x}^*$ is indeed a solution of (3.1) and (1.1).

Thus, verification of the bound in (3.15) will complete the proof. Equation (3.17) is used to express the difference between approximations of dimension n and m as,

$$\begin{aligned}(\mathbf{x}^{(n)} - \mathbf{x}^{(m)}) = (\mathbf{c}^{(n)} - \mathbf{c}^{(m)}) &+ \int_0^t A^{(n)} (\mathbf{x}^{(n)} - \mathbf{x}^{(m)}) \, d\tau \\ &+ \int_0^t (A^{(n)} - A^{(m)}) \, \mathbf{x}^{(m)} \, d\tau.\end{aligned} \tag{3.19}$$

The difference between a pair of elements then has the property

$$\begin{aligned}| x_i^{(n)} - x_i^{(m)} | \leqslant | c_i^{(n)} - c_i^{(m)} | &+ \int_0^t \sum_j | a_{ij}^{(n)} | \, | x_j^{(n)} - x_j^{(m)} | \, d\tau \\ &+ \int_0^t \sum_j | a_{ij}^{(n)} - a_{ij}^{(m)} | \, | x_j^{(m)} | \, d\tau.\end{aligned} \tag{3.20}$$

Summation over i produces the bound

$$\begin{aligned}\| \mathbf{x}^{(n)} - \mathbf{x}^{(m)} \| \leqslant \sum_{i=m+1}^{n} | c_i | &+ \sum_i \int_0^t \sum_j | a_{ij}^{(n)} | \, | x_j^{(n)} - x_j^{(m)} | \, d\tau \\ &+ \sum_i \int_0^t \sum_j | a_{ij}^{(n)} - a_{ij}^{(m)} | \, | x_j^{(m)} | \, d\tau.\end{aligned} \tag{3.21}$$

The definitions in (2.5), which have been used to achieve the simplified c_i term in (3.21), make all of the sums in (3.21) equivalent to finite sums whose orders can be interchanged freely.

The i-sums in (3.21) can be further simplified. Assuming $m \leqslant n$, the definitions in (2.5) lead to

$$\begin{aligned}\sum_i \mid a_{ij}^{(n)} - a_{ij}^{(m)} \mid &= \sum_{i=m+1}^{n} \mid a_{ij} \mid : j \leqslant m \\ &= \sum_{i=1}^{n} \mid a_{ij} \mid \leqslant \beta : m < j \leqslant n \\ &= 0; \qquad \text{otherwise}\end{aligned} \tag{3.22}$$

and

$$\sum_i \mid a_{ij}^{(n)} \mid \leqslant \beta. \tag{3.23}$$

The convergence in (2.3-c) gives the further bound, for any $\epsilon_1 > 0$,

$$\sum_{i=m+1}^{n} \mid a_{ij} \mid < \epsilon_1 , \qquad m > M_{\epsilon_1} \tag{3.24}$$

for the ($j \leqslant m$) case in (3.22). A similar bound exists for the bounded sum of $\mid c_i \mid$ in (3.21).

Substitution of the foregoing relations into (3.21) yields

$$\begin{aligned}\| \mathbf{x}^{(n)} - \mathbf{x}^{(m)} \| \leqslant \delta + \beta \int_0^t \sum_j \mid x_j^{(n)} - x_j^{(m)} \mid d\tau + \epsilon_1 \int_0^t \sum_{j=1}^{m} \mid x_j^{(m)} \mid d\tau \\ + \beta \int_0^t \sum_{j=m+1}^{n} \mid x_j^{(m)} \mid d\tau.\end{aligned} \tag{3.25}$$

Since Theorem 1 applies to each truncated system, (3.8) implies that

$$\| \mathbf{x}^{(m)} \| \leqslant \| \mathbf{c} \| e^{\beta t} \tag{3.26}$$

and, further,

$$\sum_{j=m+1}^{n} \mid x_j^{(m)} \mid \leqslant \epsilon_2 e^{\beta t}, \qquad \text{for } m > M_{\epsilon_2} \tag{3.27}$$

and any $\epsilon_2 > 0$. The last three relations combine to show that there exists an $M_{\epsilon_1,\epsilon_2,\delta}$ for every ϵ_1, ϵ_2, $\delta > 0$, such that $m > M_{\epsilon_1,\epsilon_2,\delta}$ implies

$$\| \mathbf{x}^{(n)} - \mathbf{x}^{(m)} \| \leqslant \eta + \beta \int_0^t \| \mathbf{x}^{(n)} - \mathbf{x}^{(m)} \| d\tau, \tag{3.28}$$

where

$$\eta \leqslant \delta + \epsilon_1 M e^{\beta T}/\beta + \epsilon_2 e^{\beta T}. \tag{3.29}$$

Relation (3.28) leads to [1, 2]

$$\| \mathbf{x}^{(n)} - \mathbf{x}^{(m)} \| \leqslant \eta e^{\beta t}, \tag{3.30}$$

where, from (3.29), η is arbitrarily small for all $t \in \mathscr{T}$, when m is sufficiently large. This completes the verification of (3.15) with

$$\rho = \eta e^{\beta T} \tag{3.31}$$

and, consequently, Theorem 2 is proved.

4. Example

A computational example for the case of a time invariant A matrix was given in [4] and summarized in [2]. That example arose from an approximate method for evaluating the mean-integral-squared solution of a scalar, finite order linear differential equation, when one parameter of that equation was a random variable. A time-varying version of that example, for which the present results are applicable, would have

$$\begin{gathered} \ddot{z}(t) + \rho a(t)\, \dot{z}(t) + b(t)\, z(t) = 0, \\ z(0) = 1, \qquad \dot{z}(0) = 0 \end{gathered} \tag{4.1}$$

in which ρ is a uniformly distributed random variable and $a(t)$ and $b(t)$ are, e.g., bounded for $t \in \mathscr{T}$. (The example in [2] had $a = 1$, $b = 10$.)

The desired performance measure

$$J = E\left\{\int_0^T z^2\, dt\right\} \tag{4.2}$$

can be approximated by

$$J \simeq \sum_0^N \int_0^T d_i^2\, dt \tag{4.3}$$

for large N, when the $d_i(t)$ are coefficients in the expansion

$$z(t) = \sum_{i=0}^{\infty} d_i(t)\, \varphi_i(\rho), \tag{4.4}$$

which uses $\varphi_i(\rho)$ which are orthonormal with respect to the expectation operation.

The $\mathbf{d}(t)$ vector can be expressed as the solution to an infinite system of the form (1.1) which results from substitution of (4.4) into (4.1) and use of the orthonormality conditions. The existence and approximation of solutions for that system will depend strongly on the choice of orthonormal functions. Trigonometric functions will not yield an $A(t)$ matrix satisfying the assumptions in the foregoing theorems, but Walsh functions do yield an acceptable matrix, as follows:

Comparison with [2] and [4] shows that the definitions

$$\begin{aligned} x_{2i+1} &= d_i \\ x_{2i+2} &= \dot{d}_i \end{aligned} \qquad i = 0, 1, 2,... \tag{4.5}$$

lead to an $A(t)$ matrix with the following entries:

(1) For i odd,

$$\begin{aligned} a_{ij} &= 1 \qquad \text{if} \quad j = i + 1 \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{4.6}$$

(2) For i even and j odd

$$\begin{aligned} a_{ij} &= -b(t) \qquad \text{if} \quad j = i - 1 \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{4.7}$$

(3) For i and j even

$$\begin{aligned} \frac{a_{ij}}{a(t)} &= E\{\rho\} \qquad \text{if} \qquad i = j \\ &= r\, 2^{-(p+2)} \qquad \text{if} \quad (i \dotplus j) = 2^p, \quad p = 0, 1, 2,... \\ &= 0, \qquad \text{otherwise.} \end{aligned} \tag{4.8}$$

The symbol $(i \dotplus j)$ in (4.8) means the modulo-two sum, without carrying, of the binary representations of i and j. Furthermore, r is the "width" of the uniform distribution, i.e.,

$$r^2 = 12 \operatorname{var}(\rho). \tag{4.9}$$

The matrix with elements $| a_{ij} |$ is thus symmetric with nonzero entries described by

(i) a single 1 in each odd row

(ii) some permutation of $| b(t)|$, $| a(t)| E\{\rho\}$, $r | a(t)|/4$, $r | a(t)|/8$,... in each even row. It is clear that this matrix satisfies (2.3) if $a(t)$ and $b(t)$ are uniformly bounded in $\mathscr{T}$.

Although the time varying nature of this $A(t)$ does not use the full power of the theorems presented above, the generalization of the underlying equation (4.1) permits the use of approximations like (4.3) in the important class of time-varying control system problems with random parameters. As described in [4], this method can be more efficient than Monte Carlo approximations to (4.2), and it can be generalized to cases of (4.1) having several random parameters with non-uniform distributions.

5. Remarks

The example gives some interpretation to the kind of systems which do not satisfy the more usual norm condition

$$\sum_{i,j=1}^{\infty} | a_{ij}(t)| < \infty \tag{5.1}$$

but which are solvable according to the present theorems. Expressions (4.6)–(4.8) indicate that in every row and column all entries sufficiently far from the diagonal are arbitrarily small. This decoupling of $\dot{x}_i$ from x_j for large $| i - j |$ combines with the small $x_i(0)$ for large i to produce an $\mathbf{x}(t)$ solution with bounded norm.

Weaker existence and approximation conditions for the case of time invariant A are given in [2]. In the time invariant case (2.3-a and b) become a special form of the Schur conditions [5] which make A a representation of a bounded operator on a separable Hilbert space. The results in [6] then characterize the solutions of the original differential equation.

It is straightforward to extend these results to the inhomogeneous equation

$$\dot{\mathbf{x}} = A(t)\,\mathbf{x} + \mathbf{f}(t), \tag{5.2}$$

when the forcing function $\mathbf{f}(t)$ is in $\mathscr{C}$.

Finally, it should be noted that Theorem 3 of [2] is a special case of the first theorem given here.

References

1. R. Bellman, The boundedness of solutions of infinite systems of linear differential equations, *Duke Math. J.* **14** (1947), 695–706.
2. L. Shaw, Existence and approximation of solutions to an infinite set of linear time invariant differential equations, *SIAM J. Appl. Math.* **22** (1972), 266–279.
3. W. Rudin, "Principles of Mathematical Analysis," 2nd ed., Chapters 7 and 8, McGraw-Hill, New York, 1964.

4. L. Shaw and P. Chalon, An expansion for evaluation of sensitivity with respect to a random parameter, *Automatica* **5** (1969), 265–273.
5. P. R. Halmos, "A Hilbert Space Problem Book," Problem 37, Van Nostrand, Princeton, NJ, 1967.
6. N. Dunford and J. T. Schwartz, "Linear Operators," Part I, Section VIII, 1, Interscience, New York, 1966.